RAW SEQUENCE LISTING   DATE: 08/02/2001
PATENT APPLICATION: US/09/531,266   TIME: 17:26:26

Input Set : A:\Ma258100.app
Output Set: N:\CRF3\08022001\I531266.raw

ENTERED

```
 3 <110> APPLICANT: DUNICAN, L.K.
 4        MCCORMACK, ASHLING
 5        STAPELTON, CLIONA
 6        BURKE, KEVIN
 7        MOCKEL, BETTINA
 9 <120> TITLE OF INVENTION: NEW NUCLEOTIDE SEQUENCES WHICH CODE FOR THE TAL GENE
11 <130> FILE REFERENCE: MAS/21123/258100
13 <140> CURRENT APPLICATION NUMBER: 09/531,266
14 <141> CURRENT FILING DATE: 2000-03-20
16 <150> PRIOR APPLICATION NUMBER: 60/142,915
17 <151> PRIOR FILING DATE: 1999-07-09
19 <160> NUMBER OF SEQ ID NOS: 6
21 <170> SOFTWARE: PatentIn Ver. 2.1
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 6995
25 <212> TYPE: DNA
26 <213> ORGANISM: Corynebacterium glutamicum
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (2471)..(3550)
31 <223> OTHER INFORMATION: tal-Gen
33 <400> SEQUENCE: 1
34 cacatttgaa ccacagttgg ttataaaatg ggttcaacat cactatggtt agaggtgttg 60
36 acgggtcaga ttaagcaaag actactttcg gggtagatca cctttgccaa atttgaacca 120
38 attaacctaa gtcgtagatc tgatcatcgg atctaacgaa aacgaaccaa aactttggtc 180
40 ccggtttaac ccaggaagga ttgaccacct tgacgctgtc acctgaactt caggcgctca 240
42 ctgtacgcaa ttacccctct gattggtccg atgtggacac caaggctgta gacactgttc 300
44 gtgtcctcgc tgcagacgct gtagaaaact gtggctccgg ccacccaggc accgcaatga 360
46 gcctggctcc ccttgcatac accttgtacc agcgggttat gaacgtagat ccacaggaca 420
48 ccaactgggc aggccgtgac cgcttcgttc tttcttgtgg ccactcctct ttgacccagt 480
50 acatccagct ttacttgggt ggattcggcc ttgagatgga tgacctgaag gctctgcgca 540
52 cctgggattc cttgacccca ggacaccctg agtaccgcca caccaagggc gttgagatca 600
54 ccactggccc tcttggccag ggtcttgcat ctgcagttgg tatggccatg gctgctcgtc 660
56 gtgagcgtgg cctattcgac ccaaccgctg ctgagggcga atccccattc gaccaccaca 720
58 tctacgtcat tgcttctgat ggtgacctgc aggaaggtgt cacctctgag gcatcctcca 780
60 tcgctggcac ccagcagctg ggcaacctca tcgtgttctg ggatgacaac cgcatctcca 840
62 tcgaagacaa cactgagatc gctttcaacg aggacgttgt tgctcgttac aaggcttacg 900
64 gctggcagac cattgaggtt gaggctggcg aggacgttgc agcaatcgaa gctgcagtgg 960
66 ctgaggctaa gaaggacacc aagcgaccta ccttcatccg cgttcgcacc atcatcggct 1020
68 tcccagctcc aactatgatg aacaccggtg ctgtgcacgg tgctgctctt ggcgcagctg 1080
70 aggttgcagc aaccaagact gagcttggat cgatcctga ggctcacttc gcgatcgacg 1140
72 atgaggttat cgctcacacc cgctccctcg cagagcgcgc tgcacagaag aaggctgcat 1200
74 ggcaggtcaa gttcgatgag tgggcagctg ccaaccctga gaacaaggct ctgttcgatc 1260
76 gcctgaactc ccgtgagctt ccagcgggct acgctgacga gctcccaaca tgggatgcag 1320
78 atgagaaggg cgtcgcaact cgtaaggctt ccgaggctgc acttcaggca ctgggcaaga 1380
80 cccttcctga gctgtggggc ggttccgctg acctcgcagg ttccaacaac accgtgatca 1440
82 agggctcccc ttccttcggc cctgagtcca tctccaccga gacctggtct gctgagcctt 1500
```

```
 84  acggccgtaa cctgcacttc ggtatccgtg agcacgctat gggatccatc ctcaacggca 1560
 86  tttccctcca cggtggcacc cgcccatacg gcggaacctt cctcatcttc tccgactaca 1620
 88  tgcgtcctgc agttcgtctt gcagctctca tggagaccga cgcttactac gtctggaccc 1680
 90  acgactccat cggtctgggc gaagatggcc caacccacca gcctgttgaa accttggctg 1740
 92  cactgcgcgc catcccaggt ctgtccgtcc tgcgtcctgc agatgcgaac gagaccgccc 1800
 94  aggcttgggc tgcagcactt gagtacaagg aaggccctaa gggtcttgca ctgacccgcc 1860
 96  agaacgttcc tgttctggaa ggcaccaagg agaaggctgc tgaaggcgtt cgccgcggtg 1920
 98  gctacgtcct ggttgagggt tccaaggaaa ccccagatgt gatcctcatg ggctccggct 1980
100  ccgaggttca gcttgcagtt aacgctgcga aggctctgga agctgagggc gttgcagctc 2040
102  gcgttgtttc cgttccttgc atggattggt tccaggagca ggacgcagag tacatcgagt 2100
104  ccgttctgcc tgcagctgtg accgctcgtg tgtctgttga agctggcatc gcaatgcctt 2160
106  ggtaccgctt cttgggcacc cagggccgtg ctgtctccct tgagcacttc ggtgcttctg 2220
108  cggattacca gaccctgttt gagaagttcg gcatcaccac cgatgcagtc gtggcagcgg 2280
110  ccaaggactc cattaacggt taattgccct gctgttttta gcttcaaccc ggggcaatat 2340
112  gattctccgg aattttattg ccccgggttg ttgttgttaa tcggtacaaa gggtcttaag 2400
114  cacatccctt acttgcctgc tctccttgag cacagttcaa gaacaattct tttaaggaaa 2460
116  atttagtttc atg tct cac att gat gat ctt gca cag ctc ggc act tcc        2509
117            Met Ser His Ile Asp Asp Leu Ala Gln Leu Gly Thr Ser
118             1               5                  10
120  act tgg ctc gac gac ctc tcc cgc gag cgc att act tcc ggc aat ctc        2557
121  Thr Trp Leu Asp Asp Leu Ser Arg Glu Arg Ile Thr Ser Gly Asn Leu
122      15                  20                  25
124  agc cag gtt att gag gaa aag tct gta gtc ggt gtc acc acc aac cca        2605
125  Ser Gln Val Ile Glu Glu Lys Ser Val Val Gly Val Thr Thr Asn Pro
126  30                  35                  40                  45
128  gct att ttc gca gca gca atg tcc aag ggc gat tcc tac gac gct cag        2653
129  Ala Ile Phe Ala Ala Ala Met Ser Lys Gly Asp Ser Tyr Asp Ala Gln
130                  50                  55                  60
132  atc gca gag ctc aag gcc gct ggc gca tct gtt gac cag gct gtt tac        2701
133  Ile Ala Glu Leu Lys Ala Ala Gly Ala Ser Val Asp Gln Ala Val Tyr
134                  65                  70                  75
136  gcc atg agc atc gac gac gtt cgc aat gct tgt gat ctg ttc acc ggc        2749
137  Ala Met Ser Ile Asp Asp Val Arg Asn Ala Cys Asp Leu Phe Thr Gly
138                  80                  85                  90
140  atc ttc gag tcc tcc aac ggc tac gac ggc cgc gtg tcc atc gag gtt        2797
141  Ile Phe Glu Ser Ser Asn Gly Tyr Asp Gly Arg Val Ser Ile Glu Val
142      95                  100                 105
144  gac cca cgt atc tct gct gac cgc gac gca acc ctg gct cag gcc aag        2845
145  Asp Pro Arg Ile Ser Ala Asp Arg Asp Ala Thr Leu Ala Gln Ala Lys
146  110                 115                 120                 125
148  gag ctg tgg gca aag gtt gat cgt cca aac gtc atg atc aag atc cct        2893
149  Glu Leu Trp Ala Lys Val Asp Arg Pro Asn Val Met Ile Lys Ile Pro
150                  130                 135                 140
152  gca acc cca ggt tct ttg cca gca atc acc gac gct ttg gct gag ggc        2941
153  Ala Thr Pro Gly Ser Leu Pro Ala Ile Thr Asp Ala Leu Ala Glu Gly
154                  145                 150                 155
156  atc agc gtt aac gtc acc ttg atc ttc tcc gtt gct cgc tac cgc gag        2989
157  Ile Ser Val Asn Val Thr Leu Ile Phe Ser Val Ala Arg Tyr Arg Glu
158          160                 165                 170
```

RAW SEQUENCE LISTING   DATE: 08/02/2001
PATENT APPLICATION: US/09/531,266   TIME: 17:26:26

Input Set : A:\Ma258100.app
Output Set: N:\CRF3\08022001\I531266.raw

```
160 gtc atc gct gcg ttc atc gag ggc atc aag cag gct gct gca aac ggc   3037
161 Val Ile Ala Ala Phe Ile Glu Gly Ile Lys Gln Ala Ala Ala Asn Gly
162     175                 180                 185
164 cac gac gtc tcc aag atc cac tct gtg gct tcc ttc ttc gtc tcc cgc   3085
165 His Asp Val Ser Lys Ile His Ser Val Ala Ser Phe Phe Val Ser Arg
166 190                 195                 200                 205
168 gtc gac gtt gag atc gac aag cgc ctc gag gca atc gga tcc gat gag   3133
169 Val Asp Val Glu Ile Asp Lys Arg Leu Glu Ala Ile Gly Ser Asp Glu
170                 210                 215                 220
172 gct ttg gct ctg cgc ggc aag gca ggc gtt gcc aac gct cag cgc gct   3181
173 Ala Leu Ala Leu Arg Gly Lys Ala Gly Val Ala Asn Ala Gln Arg Ala
174             225                 230                 235
176 tac gct gtg tac aag gag ctt ttc gac gcc gcc gag ctg cct gaa ggt   3229
177 Tyr Ala Val Tyr Lys Glu Leu Phe Asp Ala Ala Glu Leu Pro Glu Gly
178         240                 245                 250
180 gcc aac act cag cgc cca ctg tgg gca tcc acc ggc gtg aag aac cct   3277
181 Ala Asn Thr Gln Arg Pro Leu Trp Ala Ser Thr Gly Val Lys Asn Pro
182     255                 260                 265
184 gcg tac gct gca act ctt tac gtt tcc gag ctg gct ggt cca aac acc   3325
185 Ala Tyr Ala Ala Thr Leu Tyr Val Ser Glu Leu Ala Gly Pro Asn Thr
186 270                 275                 280                 285
188 gtc aac acc atg cca gaa ggc acc atc gac gcg gtt ctg gag cag ggc   3373
189 Val Asn Thr Met Pro Glu Gly Thr Ile Asp Ala Val Leu Glu Gln Gly
190                 290                 295                 300
192 aac ctg cac ggt gac acc ctg tcc aac tcc gcg gca gaa gct gac gct   3421
193 Asn Leu His Gly Asp Thr Leu Ser Asn Ser Ala Ala Glu Ala Asp Ala
194             305                 310                 315
196 gtg ttc tcc cag ctt gag gct ctg ggc gtt gac ttg gca gat gtc ttc   3469
197 Val Phe Ser Gln Leu Glu Ala Leu Gly Val Asp Leu Ala Asp Val Phe
198         320                 325                 330
200 cag gtc ctg gag acc gag ggt gtg gac aag ttc gtt gct tct tgg agc   3517
201 Gln Val Leu Glu Thr Glu Gly Val Asp Lys Phe Val Ala Ser Trp Ser
202     335                 340                 345
204 gaa ctg ctt gag tcc atg gaa gct cgc ctg aag tagaatcagc acgctgcatc 3570
205 Glu Leu Leu Glu Ser Met Glu Ala Arg Leu Lys
206 350                 355                 360
208 agtaacggcg acatgaaatc gaattagttc gatcttatgt ggccgttaca catctttcat 3630
210 taaagaaagg atcgtgacac taccatcgtg agcacaaaca cgaccccctc cagctggaca 3690
212 aacccactgc gcgacccgca ggataaacga ctcccccgca tcgctggccc ttccggcatg 3750
214 gtgatcttcg gtgtcactgg cgacttggct cgaaagaagc tgctccccgc catttatgat 3810
216 ctagcaaacc gcggattgct gcccccagga ttctcgttgg taggttacgg ccgccgcgaa 3870
218 tggtccaaag aagactttga aaaatacgta cgcgatgccg caagtgctgg tgctcgtacg 3930
220 gaattccgtg aaaatgtttg ggagcgcctc gccgagggta tggaatttgt tcgcggcaac 3990
222 tttgatgatg atgcagcttt cgacaacctc gctgcaacac tcaagcgcat cgacaaaacc 4050
224 cgcggcaccg ccggcaactg ggcttactac ctgtccattc caccagattc cttcacagcg 4110
226 gtctgccacc agctggagcg ttccggcatg gctgaatcca ccgaagaagc atggcgccgc 4170
228 gtgatcatcg agaagccttt cggccacaac ctcgaatccg cacacgagct caaccagctg 4230
230 gtcaacgcag tcttcccaga atcttctgtg ttccgcatcg accactattt gggcaaggaa 4290
232 acagttcaaa acatcctggc tctgcgtttt gctaaccagc tgtttgagcc actgtggaac 4350
```

```
234 tccaactacg ttgaccacgt ccagatcacc atggctgaag atattggctt gggtggacgt 4410
236 gctggttact acgacggcat cggcgcagcc cgcgacgtca tccagaacca cctgatccag 4470
238 ctcttggctc tggttgccat ggaagaacca atttctttcg tgccagcgca gctgcaggca 4530
240 gaaaagatca aggtgctctc tgcgacaaag ccgtgctacc cattggataa aacctccgct 4590
242 cgtggtcagt acgctgccgg ttggcagggc tctgagttag tcaagggact tcgcgaagaa 4650
244 gatggcttca accctgagtc caccactgag acttttgcgg cttgtacctt agagatcacg 4710
246 tctcgtcgct gggctggtgt gccgttctac ctgcgcaccg gtaagcgtct tggtcgccgt 4770
248 gttactgaga ttgccgtggt gtttaaagac gcaccacacc agcctttcga cggcgacatg 4830
250 actgtatccc ttggccaaaa cgccatcgtg attcgcgtgc agcctgatga aggtgtgctc 4890
252 atccgcttcg gttccaaggt tccaggttct gccatggaag tccgtgacgt caacatggac 4950
254 ttctcctact cagaatcctt cactgaagaa tcacctgaag catacgagcg cctcattttg 5010
256 gatgcgctgt tagatgaatc cagcctcttc cctaccaacg aggaagtgga actgagctgg 5070
258 aagattctgg atccaattct tgaagcatgg gatgccgatg gagaaccaga ggattaccca 5130
260 gcgggtacgt ggggtccaaa gagcgctgat gaaatgcttt cccgcaacgg tcacacctgg 5190
262 cgcaggccat aatttagggg caaaaaatga tctttgaact tccggatacc accacccagc 5250
264 aaatttccaa gaccctaact cgactgcgtg aatcgggcac ccaggtcacc accggccgag 5310
266 tgctcaccct catcgtggtc actgactccg aaagcgatgt cgctgcagtt accgagtcca 5370
268 ccaatgaagc ctcgcgcgag cacccatctc gcgtgatcat tttggtggtt ggcgataaaa 5430
270 ctgcagaaaa caaagttgac gcagaagtcc gtatcggtgg cgacgctggt gcttccgaga 5490
272 tgatcatcat gcatctcaac ggacctgtcg ctgacaagct ccagtatgtc gtcacaccac 5550
274 tgttgcttcc tgacaccccc atcgttgctt ggtggccagg tgaatcacca aagaatcctt 5610
276 cccaggaccc aattggacgc atcgcacaac gacgcatcac tgatgctttg tacgaccgtg 5670
278 atgacgcact agaagatcgt gttgagaact atcacccagg tgataccgac atgacgtggg 5730
280 cgcgccttac ccagtggcgg ggacttgttg cctcctcatt ggatcaccca ccacacagcg 5790
282 aaatcacttc cgtgaggctg accggtgcaa gcggcagtac ctcggtggat ttggctgcag 5850
284 gctggttggc gcggaggctg aaagtgcctg tgatccgcga ggtgacagat gctcccaccg 5910
286 tgccaaccga tgagtttggt actccactgc tggctatcca gcgcctggag atcgttcgca 5970
288 ccaccggctc gatcatcatc accatctatg acgctcatac ccttcaggta gagatgccgg 6030
290 aatccggcaa tgccccatcg ctggtggcta ttggtcgtcg aagtgagtcc gactgcttgt 6090
292 ctgaggagct tcgccacatg gatccagatt tgggctacca gcacgcacta tccggcttgt 6150
294 ccagcgtcaa gctggaaacc gtctaaggag aaatacaaca ctatggttga tgtagtacgc 6210
296 gcacgcgata ctgaagattt ggttgcacag gctgcctcca aattcattga ggttgttgaa 6270
298 gcagcaactg ccaataatgg caccgcacag gtagtgctca ccggtggtgg cgccggcatc 6330
300 aagttgctgg aaaagctcag cgttgatgcg gctgaccttg cctgggatcg cattcatgtg 6390
302 ttcttcggcg atgagcgcaa tgtccctgtc agtgattctg agtccaatga gggccaggct 6450
304 cgtgaggcac tgttgtccaa ggtttctatc cctgaagcca acattcacgg atatggtctc 6510
306 ggcgacgtag atcttgcaga ggcagcccgc gcttacgaag ctgtgttgga tgaattcgca 6570
308 ccaaacggct ttgatcttca cctgctcggc atgggtggcg aaggccatat caactccctg 6630
310 ttccctcaca ccgatgcagt caaggaatcc tccgcaaagg tcatcgcggt gtttgattcc 6690
312 cctaagcctc cttcagagcg tgcaactcta acccttcctg cggttcactc cgcaaagcgc 6750
314 gtgtggttgc tggtttctgg tgcggagaag gctgaggcag ctgcggcgat cgtcaacggt 6810
316 gagcctgctg ttgagtggcc tgctgctgga gctaccggat ctgaggaaac ggtattgttc 6870
318 ttggctgatg atgctgcagg aaatctctaa gcagcgccag ctctaacaag aagctttaac 6930
320 aagaagctct aacgaaaagc actaacaaac taatccgggt gcgaaccttc atctgaatcg 6990
322 atgga                                                               6995
325 <210> SEQ ID NO: 2
326 <211> LENGTH: 360
327 <212> TYPE: PRT
328 <213> ORGANISM: Corynebacterium glutamicum
```

RAW SEQUENCE LISTING   DATE: 08/02/2001
PATENT APPLICATION: US/09/531,266   TIME: 17:26:26

Input Set : **A:\Ma258100.app**
Output Set: **N:\CRF3\08022001\I531266.raw**

```
330 <400> SEQUENCE: 2
331 Met Ser His Ile Asp Asp Leu Ala Gln Leu Gly Thr Ser Thr Trp Leu
332  1            5                   10                  15
334 Asp Asp Leu Ser Arg Glu Arg Ile Thr Ser Gly Asn Leu Ser Gln Val
335             20                  25                  30
337 Ile Glu Glu Lys Ser Val Val Gly Val Thr Thr Asn Pro Ala Ile Phe
338             35                  40                  45
340 Ala Ala Ala Met Ser Lys Gly Asp Ser Tyr Asp Ala Gln Ile Ala Glu
341      50                  55                  60
343 Leu Lys Ala Ala Gly Ala Ser Val Asp Gln Ala Val Tyr Ala Met Ser
344  65                  70                  75                  80
346 Ile Asp Asp Val Arg Asn Ala Cys Asp Leu Phe Thr Gly Ile Phe Glu
347                  85                  90                  95
349 Ser Ser Asn Gly Tyr Asp Gly Arg Val Ser Ile Glu Val Asp Pro Arg
350              100                 105                 110
352 Ile Ser Ala Asp Arg Asp Ala Thr Leu Ala Gln Ala Lys Glu Leu Trp
353          115                 120                 125
355 Ala Lys Val Asp Arg Pro Asn Val Met Ile Lys Ile Pro Ala Thr Pro
356      130                 135                 140
358 Gly Ser Leu Pro Ala Ile Thr Asp Ala Leu Ala Glu Gly Ile Ser Val
359 145                 150                 155                 160
361 Asn Val Thr Leu Ile Phe Ser Val Ala Arg Tyr Arg Glu Val Ile Ala
362              165                 170                 175
364 Ala Phe Ile Glu Gly Ile Lys Gln Ala Ala Ala Asn Gly His Asp Val
365              180                 185                 190
367 Ser Lys Ile His Ser Val Ala Ser Phe Phe Val Ser Arg Val Asp Val
368          195                 200                 205
370 Glu Ile Asp Lys Arg Leu Glu Ala Ile Gly Ser Asp Glu Ala Leu Ala
371      210                 215                 220
373 Leu Arg Gly Lys Ala Gly Val Ala Asn Ala Gln Arg Ala Tyr Ala Val
374 225                 230                 235                 240
376 Tyr Lys Glu Leu Phe Asp Ala Ala Glu Leu Pro Glu Gly Ala Asn Thr
377              245                 250                 255
379 Gln Arg Pro Leu Trp Ala Ser Thr Gly Val Lys Asn Pro Ala Tyr Ala
380              260                 265                 270
382 Ala Thr Leu Tyr Val Ser Glu Leu Ala Gly Pro Asn Thr Val Asn Thr
383          275                 280                 285
385 Met Pro Glu Gly Thr Ile Asp Ala Val Leu Glu Gln Gly Asn Leu His
386      290                 295                 300
388 Gly Asp Thr Leu Ser Asn Ser Ala Ala Glu Ala Asp Ala Val Phe Ser
389 305              310                 315                 320
391 Gln Leu Glu Ala Leu Gly Val Asp Leu Ala Asp Val Phe Gln Val Leu
392              325                 330                 335
394 Glu Thr Glu Gly Val Asp Lys Phe Val Ala Ser Trp Ser Glu Leu Leu
395              340                 345                 350
397 Glu Ser Met Glu Ala Arg Leu Lys
398          355                 360
402 <210> SEQ ID NO: 3
403 <211> LENGTH: 1083
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/531,266

DATE: 08/02/2001
TIME: 17:26:27

Input Set : A:\Ma258100.app
Output Set: N:\CRF3\08022001\I531266.raw